DOCUMENT RESUME

ED 452 494 CS 014 354

| | |
|---|---|
| AUTHOR | Grossen, Bonnie; Coulter, Gail; Ruggles, Barbara |
| TITLE | Reading Recovery: An Evaluation of Benefits and Costs. Executive Summary. |
| PUB DATE | 1997-00-00 |
| NOTE | 23p.; Cover page varies. |
| AVAILABLE FROM | For full text: http://www.darkwing.uoregon.edu/~bgrossen/rr.htm. |
| PUB TYPE | Information Analyses (070) -- Reports - Evaluative (142) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | *Early Intervention; Grade 1; Instructional Effectiveness; Literature Reviews; Primary Education; *Program Effectiveness; Program Evaluation; *Reading Difficulties; *Reading Instruction; Remedial Reading |
| IDENTIFIERS | *Reading Recovery Projects; Research Summaries |

ABSTRACT

In Reading Recovery, program-trained teachers provide one-to-one tutoring in 30-minute daily sessions to the lowest 10 to 20% of a first-grade class who have the prerequisite skills for Reading Recovery. Reading Recovery advocates claim that the program brings the lowest performing children up to the average level of their local class by the end of first grade within 60 lessons, or 12 weeks. When students reach this goal they are "discontinued" from the program. Because of Reading Recovery's increasing popularity and its expense, more independent evaluators are raising questions and reviewing the research that is cited to support claims regarding its effectiveness. This paper presents a summary of these reviews and other studies evaluating the impact of Reading Recovery--these findings should be considered in deciding whether to adopt, expand, or terminate Reading Recovery programs. The paper concludes with the following points: (1) the Reading Recovery data reporting system is flawed; (2) the standards for successful completion of Reading Recovery is not equitable; (3) Reading Recovery does not raise overall school achievement levels; (4) far fewer students than claimed actually benefit from Reading Recovery; (5) children who are not expected to be successful are removed from the program and from the calculation of the success rate; (6) Reading Recovery does not reduce the need for other compensatory reading services; and (7) research-based alternative interventions are more effective than Reading Recovery. (Contains 4 figures, a table, and 53 references.) (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 452 494

# Reading Recovery: An Evaluation of Benefits and Costs.

by

Bonnie Grossen, Gail Coulter, and Barbara Ruggles

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

B. Grossen

CS 014 354

# READING RECOVERY:

# AN EVALUATION OF BENEFITS AND COSTS

**Bonnie Grossen and Gail Coulter**

**University of Oregon**

**Barbara Ruggles**

**Beacon Hill Elementary, Park Forest, Illinois**

## Executive Summary

Reading Recovery is being widely adopted in North America:

*"Reading Recovery sites operated in four Canadian provinces, 48 U.S. States, and the District of Columbia. Approximately 60,000 North American children were served by Reading Recovery educators during the 1993-94 school year. In California alone, more than 500 school districts served approximately 5000 children."* (Schwartz & Klein, 1996)

Many believe Reading Recovery to be the best available program for preventing reading failure. Reading Recovery was developed in the 1970s by Dr. Marie Clay, a New Zealand educator, to deal with the reading failure occurring there. It was introduced in the United States through the Ohio State University in 1984 by Dr. Gay Su Pinnell and Dr. Charlotte Huck. Gay Sue Pinnell, Diane Deford, and Carol A. Lyons are directors of the National Reading Recovery Center at Ohio State in the U.S.

In Reading Recovery, program-trained teachers provide one-to-one tutoring in 30-minute daily sessions to the lowest 10 to 20% of a first-grade class who have the prerequisite skills for Reading Recovery. Reading Recovery advocates claim that the program brings the lowest performing children up to the average level of their local class by the end of first grade within 60 lessons, or 12 weeks. When students reach this goal they are "discontinued" from the Reading Recovery program, at which time the Reading Recovery teacher can take another student into the 30-minute slot. Each Reading Recovery-trained teacher, working a half-day with Reading Recovery, is expected to be able to tutor 8 students in one year, though actual figures from the national data set indicate that the average number of students per teacher is much lower-5.5, or 11 students for a full-time equivalent teacher, according to Hiebert (1994).

Because of Reading Recovery's increasing popularity, and its expense, more independent evaluators are raising questions and reviewing the research that is cited to support claims regarding its effectiveness. Following is a summary of the findings of these reviews and other studies evaluating the impact of Reading Recovery. These findings should be considered in deciding whether to adopt, expand, or terminate Reading Recovery programs.

***The Reading Recovery data reporting system is flawed.***

The in-house Reading Recovery evaluation system results in considerable bias in the data collected through that system. Persons responsible for success collect the data on success. Without explanation, about half the data on children eligible for Reading Recovery are omitted from final analyses (Shanahan & Barr, 1995). In addition, the measures used to evaluate Reading Recovery *(Clay Diagnostic Measures)* emphasize tasks that align with the specific strategies taught in Reading Recovery (Center, Wheldall, & Freeman, 1992; Wasik & Slavin, 1993). For example, the children are taught to use context to predict words rather than sounding them out. The reading measure uses predictable text, rather than text that uses authentic, natural language patterns. Children who have learned the prediction strategies of Reading Recovery will score better reading predictable text than they will reading authentic text. Because of the close alignment of the measures with the strategies taught in Reading Recovery, the results of an evaluation using these measures are biased in favor of Reading Recovery.

***The standard for successful completion of Reading Recovery is not equitable.***

Reading Recovery's goal to bring the lowest pupils to the average level of their class, falls short of a more equitable ***standard*** level, such as the national average. The average level of performance of a class of children from low income areas is about the 20th percentile on national norm-referenced measures. ("Grade level" is the 50th percentile.) In inner-city schools where so many students do not learn to read, only a few students can be served with Reading Recovery. Some of the lowest children will be brought up to only the 20th percentile and many children performing below the 20th percentile will not be served. As a statewide intervention Reading Recovery would result in allocating the same resources to the goal of raising a few children in a low income school to the 20th percentile that it would allocate to a high income school raising children scoring below the 80th percentile to the 80th percentile. This inequity raises constitutional issues because it impacts minority children, who are overrepresented in low income schools. Average first-grade children are more likely to be nonreaders in low income schools.

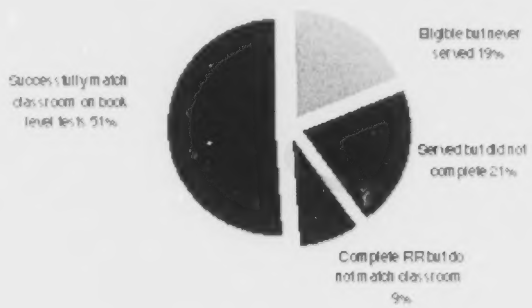
***Reading Recovery does not raise overall school achievement levels.***

If a school's goal is to raise the overall level of reading performance, Reading Recovery is not the appropriate intervention to choose. Overall school achievement scores are not improved with the use of Reading Recovery (Hiebert, 1994). Both Reading Recovery advocates and critics agree on this point (Hiebert, 1994; Pinnell & Lyons, 1995)

***Far fewer students than claimed actually benefit from Reading Recovery.***

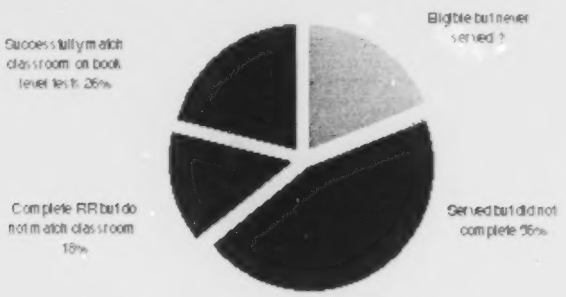
Analyses reporting that 75 to 85% of the children in Reading Recovery are successful are misleading because (a) nearly half the data are systematically omitted from the analyses (Shanahan & Barr, 1995), and (b) successful does not mean the children are readers. Successful is defined as being able to read text level measures at the average level of the child's class. Various independent evaluations have accounted for the missing data (Battelle, 1995; Shanahan & Barr, 1995). Figures 1 and 2 present these findings in graphic form. In both figures the black areas represent the proportion of children who were served in Reading Recovery and the grey areas represent an estimate of the children who were eligible but were not served. Figure 1 shows the national Reading Recovery data that were gathered through the in-house data collection system. Figure 2 shows the Columbus, Ohio data that were gathered by an independent evaluator (Pollock, 1996) and reported as percentages of children served (shown in black).

**Figure 1. National Reading Recovery**


Eligible but never served 19%
Successfully match classroom on book level tests 31%
Served but did not complete 21%
Complete RR but do not match classroom 9%

BEST COPY AVAILABLE

**Figure 2. Outcomes in Columbus, Ohio 1**


Eligible but never served 19
Served but did not complete 36%
Complete RR but do not match classroom 18%
Successfully match classroom on book level tests 26%

Both evaluations omitted the number of children who are eligible but never served-often because they lacked prerequisite skills or were already identified for special education. Battelle (1995) is the only source that has reported this number (19%) in an evaluation of Ohio's Reading Recovery program. Battelle's figure is used in both figures (shown in grey). Children served but who do not complete Reading Recovery include children who are removed because they do not make adequate progress. These children are *not* counted in the calculation of Reading Recovery success rates. Excluding eligible children who are never served and served children who do not complete the program for various reasons inflates the success rate. In reality, the success rate describes how accurately the Reading Recovery teacher was able to predict which students would be able to match the classroom average on the Clay Diagnostic Measures upon completion of the program. Those the teacher predicted would not succeed, s/he should have removed from the program prior to completion.

***Reading Recovery does not reduce the need for other compensatory reading services.***

Reading Recovery does not eliminate the need for Title I. Pollack (1996) reports that in Columbus, Ohio, in the 1995-6 school year, only 14.7% of the children who completed the program reached national norms, and 81% of those completing the program still remained eligible for Title I services. When all eligible children are included in the calculation, only an estimated 6.5% reached national norms and 92% continued to be eligible for Title I after Reading Recovery was implemented. (Those who are never served or who do not complete Reading Recovery remain eligible for Title I services also.) Even among the smaller portion of children counted as successful over an eight-year period by Reading Recovery standards, 31% were still eligible for Title I services (Pollock, 1994).

Reading Recovery does not eliminate the need for special education. Six or 7% of the children who are served are referred to special education (Shanahan & Barr, 1995). Wake County Public School System (WCPPS) in North Carolina found that Reading Recovery students, *"compared to a control group, were just as likely to be retained, placed in special education, or served in [Title] I a year later"* (1995, p. ii). Reading Recovery does not serve the lowest performing children. The average entry level percentile score of children who complete Reading Recovery is 34.5 (Hiebert, 1994).

***Children successful in Reading Recovery are often not successful later.***

Other research has documented that children who complete Reading Recovery and return to the class do not continue to learn at the same rate as average children in the class, but seem to immediately begin falling behind again (DeFord, Pinnell, Lyons, and Place, 1990; Glynn, Crooks, Bethune, Ballard, and Smith, 1989; Shanahan & Barr, 1995). The learning rate of returned Reading Recovery children was slower than that of other low-achieving children (Glynn, Crooks, Bethune, Ballard, & Smith, 1989).

***Research-based alternative interventions are more effective than Reading Recovery.***

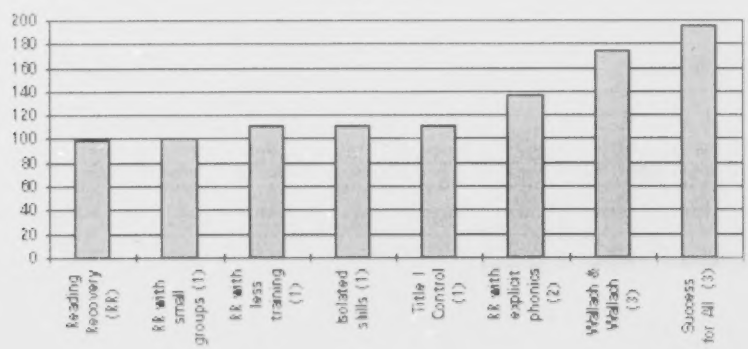
Independent evaluations have compared Reading Recovery with other common compensatory programs (Battelle, 1995; Fincher, 1991; WCPPS, 1995) and found no advantage for Reading Recovery on measures using authentic text (the natural text used in the reading comprehension passages of

standardized measures). One frequently cited study found Reading Recovery superior to other interventions (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994). Pinnell et al. compared specific variations of Reading Recovery and found approximately equal results regardless of whether the teachers had less training or the instruction was delivered in groups of four. Rasinski (1995) found serious methodological flaws in the Pinnell et al. study. He adjusted the scores to hold instructional time equal and found that the effect of Reading Recovery was at best only equivalent to the other treatments on measures of authentic text (Gates-McGinitie). Fincher (1991) compared the performance of children in Reading Recovery with that of children in other compensatory programs in Canton City Schools, Ohio, over a five-year period and found that common Title I programs resulted in better performance on measures using authentic text and other standardized measures.

*"Teaching Assistants with almost no training and minimal teaching materials with which to teach and working in less than desirable conditions, outperformed the Reading Recovery teachers when their students' overall achievement was compared. Also, Reading Recovery teachers, when their Reading Recovery students are compared with their Chapter I students, tend to get better results with the regular Chapter I program than with Reading Recovery. This has been the case every year since 1985-86, the year Reading Recovery was implemented in Canton."* (Fincher, 1991)

Research shows that explicit instruction in phonemic awareness beginning in kindergarten followed by explicit systematic instruction in phonics combined with extensive practice reading decodable text are emerging as important factors in the effective treatment of reading disabilities. Iversen and Tunmer (1993) added a component of systematic phonics to Reading Recovery. Reading Recovery with systematic phonics was 37% more efficient. Wasik and Slavin (1993) compared the relative effect sizes achieved by five treatments for reading problems. Reading Recovery was not nearly as effective as two programs that provided explicit systematic phonics with extensive practice reading decodable text (the Success for All and Wallach and Wallach programs). Decodable text is quite different from the predictable text used for practice in Reading Recovery.

**Figure 3. Relative Effect Size of Various Compensatory Programs**

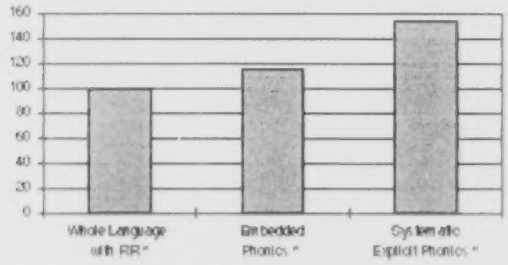



(1) Rasinski,1995
(2) Iversen & Tunmer, 1993
(3) Wasik & Slavin, 1993

Very recently the research program of the National Institute of Child Health and Human Development (Foorman, Francis, Beeler, Winikates, and Fletcher, in press) has found that changing the regular classroom program from whole language to incorporate explicit instruction in phonemic awareness and systematic phonics with decodable text is more effective than tutorial programs in reducing the occurrence of reading disabilities. Foorman et al. (in press) compared (a) whole language combined with an unlicensed Reading Recovery model, (b) embedded phonics, a semi-systematic program, and (c) explicit phonemic awareness with systematic explicit phonics and decodable text. All these treatments were delivered in the regular classroom. The explicit systematic phonic approach was more than 1 ½ times as effective in preventing reading disabilities as whole language combined with the unlicensed

Reading Recovery program (see Figure 4).

**Figure 4. Relative Effect Size of Class Interventions**




* Foorman, Francis, Beeler, Winikates, and Fletcher, in press

***Reading Recovery is extremely expensive and does not save other costs.***

Thirty hours of instruction for one child in Reading Recovery costs more than a full year of schooling for the child. Reading Recovery advocates argue that even when the highest cost estimates are used, the expense is cheap because the multi-year educational costs of special education and Title I are saved, as are the social costs of letting children fail to learn to read. However, best estimates indicate that approximately 90% of the children eligible for Reading Recovery services continue to need other compensatory services. Other alternative models are more effective. Many of these models are classwide and actually cost much less, affect more students, produce higher performance, and, most importantly, change school and classroom practices so that the need for costly after-the-fact interventions are minimized. For the cost of one year of Reading Recovery in a school, class sizes could be reduced and the whole school's early literacy program could be redesigned. By adopting research-supported best practices and whole-school change, schools could significantly increase the number of students who can read authentic grade level text. Installing a more effective school-wide program is a one-time-only investment while Reading Recovery requires the same level of investment year after year.

## READING RECOVERY:

## AN EVALUATION OF BENEFITS AND COSTS

### Research Methodology

***Does the Reading Recovery research design allow conclusions regarding program effectiveness?***

Reading Recovery includes not only an instructional program but also its own evaluation system that aligns with the program. Most of the data cited regarding the effectiveness of Reading Recovery are gathered through the Reading Recovery evaluation system. This system uses a unique pre- posttest research design and the Clay Diagnostic measures to assess student performance, both designed by Marie Clay, the Reading Recovery program developer. Close alignment of the research design, the measures, and the program, along with data collection procedures that are controlled within the Reading Recovery implementation system creates an increased potential for bias in the results of an evaluation. Because most of the data available regarding the success of Reading Recovery come from its own evaluation system, the research design and the measures used in this system are discussed first.

**The Reading Recovery Research Design**

The Reading Recovery research design is not adequate for concluding that Reading Recovery is a

superior intervention. The research design specifies that comparison groups be selected at random from the Reading Recovery students' respective classrooms. The measures are administered to these children who then represent the average for that particular first grade class. Two types of data are used to compare the performance of the Reading Recovery children with that of the comparison group. First, the achievement of the comparison group is used to establish a band of achievement. The "band" is a half standard deviation above and below the mean in each of the areas taught to the Reading Recovery students and measured by the Clay Diagnostic measures. If a Reading Recovery student's scores end up within this band, then the child is considered successful and is "discontinued" (Fincher, 1991). Secondly, the data are analyzed to compare the pre- post-gains made by the Reading Recovery children with the comparison group to see if the children in Reading Recovery gained at a faster than normal rate while in Reading Recovery.

This design is similar to that used in curriculum-based measurement (CBM), which is widely used for special education decision-making. However, there are two important differences: (a) in CBM the measures sample the class's curriculum to determine when a child is ready to be returned to the classroom; in Reading Recovery the measures sample the pull-out curriculum, and (b) in CBM conclusions are made regarding individual students so local norms are appropriate; in Reading Recovery local norms are used to evaluate the effectiveness of Reading Recovery for a whole group of students. Local norms are not appropriate for program evaluation without reference to national norms, because local norms are highly variable, and there is no way to know whether an alternative program may have been more effective without an equivalent comparison group.

**The Measures**

The Clay Diagnostic measures are used in the Reading Recovery evaluation system. Results obtained with these measures are somewhat misleading for two reasons: (a) content bias, and (b) unequal intervals between levels.

***Content bias.*** As Wasik and Slavin (1993) and Center, Wheldall, and Freeman (1992) point out, the Clay Diagnostic measures sample the specific skills taught in the Reading Recovery program. *"There is an articulation between the Reading Recovery program and the measures used to evaluate the program, suggesting that what is taught is what is measured"* (Wasik & Slavin, 1993, p. 187). This is particularly true in the lower levels of the program, where assessments emphasize less authentic reading tasks and skills that are unique to Reading Recovery. The comparison children may have no experience with the kinds of tasks evaluated by these measures, while Reading Recovery children have extensive experience. Comparisons on these measures are likely to exaggerate the amount of learning for Reading Recovery children.

The primary evaluation tool in the Clay Diagnostic measures is the book-level measure, which is used to determine where a child places in the 20 levels of the instructional program booklets. It is the only measure in the battery that requires the children to read connected text. Though the text is connected, it is not authentic text. It is "predictable" text, where pictures and repetitive sentence patterns prompt the reader. Predictability is strongest at the lowest level and is gradually reduced as children progress into the higher levels. At the final 20$^{th}$ level the text is least predictable, but it still has predictable features limiting its authenticity. Children generally do not reach the 20$^{th}$ level before they are discontinued, since they only need to reach the class average to be returned to their classroom.

The national Reading Recovery data indicate that the average level at completion of Reading Recovery is only level 10 (Shanahan & Barr, 1995). At level 10, the texts are still very predictable so the children can read words without looking closely at them. The children rely more on the contextual clues, the illustrations and the repeated sentence patterns in the text. Children who use these contextual strategies to read are more likely to be successful in predictable text than in authentic text. Consequently, children from Reading Recovery may not read authentic text very well at all when they are returned to the classroom as "successful."

Stanovich and Stanovich (1995) report that many studies have found that authentic text is not very

predictable:

*"It is often incorrectly assumed that predicting upcoming words in sentences is a relatively easy and highly accurate activity. Actually, many different empirical studies have indicated that naturalistic text is not that predictable. Alford (1980) found that for a set of moderately long expository passages of text, subjects needed an average of more than four guesses to correctly anticipate upcoming words in the passage (the method of scoring actually makes this a considerable underestimate). Across a variety of subject populations and texts, a reader's probability of predicting the next word in a passage is usually between .20 and .35 (Aborn, Rubenstein, & Sterling, 1959; Gough, 1983; Miller & Coleman, 1967; Perfetti, Goldman, & Hogaboam, 1979; Rubenstein & Aborn, 1958). Indeed, as Gough (1983) has shown, the figure is highest for function words, and is often quite low for the very words in the passage that carry the most information content."* (p. 90)

If authentic text is not very predictable, then children who read well in predictable text may not necessarily read well in authentic text. The strategies they have learned for reading may not generalize to real reading. These are important research questions that will be discussed in the review of empirical findings below.

***Unequal intervals.*** Center, Wheldall, and Freeman (1992) point out that not only are the book-level measures biased to show positive results for the prediction strategies taught in Reading Recovery, but they are also biased to show greater growth on pre-post comparisons for lower performing children:

*"Data reported by Glynn et al. (1989) indicated that the relationship between the amount of instruction and reading performance was not linear with respect to text level. Over a given time period, the average increase in text level was greater for the lower level texts than for the higher level texts (Iversen & Tunmer, in press)."* (Center, Wheldall, & Freeman, 1992, p. 271)

Because the intervals between levels are smaller at the lower levels, greater gains for poorer readers in Reading Recovery may be spurious (Center, Wheldall, & Freeman, 1992). A lower-performing Reading Recovery child learns much less to move from level 1 to level 2 than an average performing child must learn to move from level 11 to 12. Even though these intervals are not equal, a Reading Recovery evaluation would interpret these as equal gains.

**Data Collection Procedures**

The data collection process is not objective or independent. Those who collect and collate Reading Recovery success data have high stakes invested in the success of Reading Recovery. Reading Recovery teachers collect and collate success data for the children they teach. **The supervisor uses the success data collected by the teacher to evaluate the same teacher.** The supervisors then collate the data from the teachers they supervise to submit to their respective university training centers who use the data to evaluate the supervisors' performance. The national Reading Recovery directors at Ohio State University have collated the data from all the university training centers in reports to the National Diffusion Network, which has validated Reading Recovery as an effective research-based program based on these data.

Two aspects of the data collection procedures result in misleading calculations of success rates:

1. Children that the Reading Recovery teacher judges as not likely to be successful are not taken into the program. This judgment is based on entry level assessment, on a child's performance in the pre-program phase of "roaming around the known," or on other unspecified indicators. These excluded children are not counted among the children "served" by Reading Recovery, and, therefore, are not included in the calculation of the success rate.

2. Among children served some do not complete all 60 lessons. These children are also not counted in the success rate calculation. Sometimes these children are removed from Reading Recovery on the grounds that the program is not appropriate for them. Six to 7% are referred to special education (Shanahan & Barr, 1995). The others are generally referred to the Title I program. Some children fail to

complete the program because the year ends before they are finished and Reading Recovery is only for first-time first graders. (Retained first graders are not eligible.)

The needs of these children must either remain unserved or must be served by other compensatory programs. By repeatedly reducing the number of children counted in the total, the success rates reported for Reading Recovery are inflated.

**Implications for This Review**

Because of the high levels of publicity that have been given the in-house evaluations of Reading Recovery and the built-in biases contained in the in-house evaluation system, the following review emphasizes the findings of independent evaluations. Two types of independent evaluations are available: (a) independent reviews of the data gathered through the Reading Recovery evaluation system, and (b) the results obtained on independent measures of children's ability to read authentic text. Even the independent reviews rely heavily on the data collected through the Reading Recovery in-house evaluation system simply because the other data are extremely limited. Hiebert (1994) and Shanahan and Barr (1995), for example, did not collect their own data on Reading Recovery but critiqued the analyses and conclusions made from data collected by other researchers, for the most part, data from the Reading Recovery evaluation system.

Very little data have been gathered comparing Reading Recovery with alternative programs or a control group. Most of these comparative studies have also been conducted by the Reading Recovery leaders, Pinnell, Huck, Lyons, and others who direct the Reading Recovery program nationally from Ohio State University. Independent evaluators include Battelle (1995) for the Ohio Department of Education, the Wake County Public School System in North Carolina (1995), Pollock (1996) for the Columbus Ohio Public Schools, and Fincher (1991) for the Canton Ohio Public Schools. These evaluators may have no stake in Reading Recovery but often include the data collected by the in-house system in their evaluations. The independence of the evaluators makes these analyses important.

**Recent Independent Evaluations of the Reading Recovery Research Design**

The North Central Regional Educational Laboratory (NCREL), a federally supported educational laboratory responsible for interpreting educational research for the midwestern states, hired two scholars to review all the existing empirical research regarding the effectiveness of Reading Recovery. NCREL selected Rebecca Barr and Timothy Shanahan because they had articulated two opposing viewpoints regarding Reading Recovery. Barr is a noted advocate for Reading Recovery, having served on various boards for the Reading Recovery effort and as sponsor for her university's Reading Recovery training program. On the other hand, Shanahan is a noted critic of Reading Recovery, having written the first published critique of Reading Recovery research (1987).

By considering the perspectives of both sides, Shanahan and Barr's 1995 review *Reading Recovery: An Independent Evaluation of the Effects of an Early Instructional Intervention for At Risk Learners*, provides perhaps the most thorough analysis available of the data collected through the Reading Recovery evaluation system described above. Their basic finding was:

***"...that Reading Recovery leads to learning....It is less effective and more costly than has been claimed, and does not lead to systematic changes in classroom instruction, making it difficult to maintain learning gains. This is discouraging given program claims and its great expense"*** (p.1).

In addition to finding the reports of success misleading, Shanahan and Barr (1995) found unorthodox research procedures. For example, most of the Reading Recovery system data was located in unpublished technical reports produced by Reading Recovery leaders, Pinnell, Huck and others, at Ohio State University and had not undergone the peer review that is necessary to publish a study in scientific journals. A recent example of an unpublished technical report is an evaluation of Reading Recovery in Arkansas distributed by the Southern Regional Education Board (1996). This evaluation presents conclusions formed from data collected through the in-house Reading Recovery system and was not reviewed by an independent party.

All the studies Shanahan and Barr located suffered methodological problems: *"We found no studies of Reading Recovery that did not suffer from serious methodological or reporting flaws-published or not."* (1995, p. 961) Shanahan and Barr identified three types of problems in the Reading Recovery pre-post design, which would lead to exaggerated success rates:

*[The reported learning gain] most certainly is an overestimate of typical amounts of learning from Reading Recovery for several reasons: (a) test score improvements not linked to learning are likely to occur when students with extreme scores are selected for participation; (b) normal development and learning gains typical of young children can be due to other sources of growth and education; and (c) there is systematic omission of children who are not having success in Reading Recovery.* (p. 969)

The first two problems would be removed if equivalent groups of children were used as experimental controls. The systematic omission of data is a more serious problem because among those omitted are children the Reading Recovery teachers identify as ones who are not progressing well. Children who are not successful are intentionally dropped before completing the entire program. The reports then do not reflect how well Reading Recovery serves the entire population it claims to serve, nor do they provide information regarding overall class effects or school effects. Consequently, the success rates cannot be used to evaluate the effectiveness of Reading Recovery.

*Probably the most serious flaw in Reading Recovery research has to do with who is included in the experimental sample. In some analyses, only discontinued students were examined, making the program appear more effective than it really is. In most of the studies, students were omitted from analysis because of serious learning problems, poor school attendance, or other similar difficulties. These omissions were often made without mention. It is impossible to provide a valid estimate of the effects of Reading Recovery unless all children who start the program are included in the eventual analysis....Unfortunately, even two of the more sophisticated studies (Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994) that we analysed have lost as much as half of their data, without any empirical estimate or control of the effects of these missing data.* (p. 991-2)

*"The Ohio State programs have routinely collected information on those who are dropped for various reasons, but these data have not been taken account of in their studies or technical reports, nor have they been available to the public. Depending on the proportion of participants omitted in this fashion, this creates a substantial bias in favor of Reading Recovery gains, and there is no sound way to adjust the scores that are reported on this basis."* (Shanahan & Barr, 1995, p. 966)

For these reasons, Shanahan and Barr (1995) found it impossible to use standard research procedures:

*"Overall, our consideration of the existing research and evaluation studies of Reading Recovery is largely qualitative. It would be difficult or impossible to conduct a thorough empirical examination of this work using procedures such as meta-analysis because there are so few studies, and those that exist usually provide insufficient information to make such analysis appropriate.* (1995, p. 961)

## Equity

***Is the standard for successful completion of Reading Recovery equitable?***

**The standard for successful completion is not equitable.** Reading Recovery systematically results in lower expectations for children in lower achieving schools by bringing a child to only the average level of the other first-grade children in the child's class or school and not to a uniform national standard. The average level of performance of children in low income areas is approximately the 20th percentile, while the average level of children in higher income areas may be around the 80th percentile. To bring each child to the average level of the first-grade children in the child's local school leads to inequity. Children reading at only the 20th percentile in first grade are generally nonreaders and are likely to remain

unsuccessful in school, while children reading at the 80th percentile in first grade are likely to be readers.

***The relative notion of reading disability is problematic in America's poorest schools. In these schools accomplishing an instructional setting [class or school] average can mean returning children to the classroom with reading levels in the bottom 15 to 25% of the national distribution--a level of performance that, even if maintained, makes it likely that the child will not complete school successfully.*** (Shanahan & Barr, 1995, p. 995)

As a statewide intervention, Reading Recovery would not be able to make readers out of all the lowest performing children. First, Reading Recovery does not make children readers. "Success" in Reading Recovery rarely means the children read. Second, low income schools with heavy concentrations of nonreaders need a school-wide intervention, not a tutorial that works with 10 to 20% of the children. Reading Recovery is least effective in the lowest performing schools because of the high proportion of students who are reading below the 25th percentile. In this case Reading Recovery can serve only a small percentage of students who are significantly behind. A large number of children who need services are left without assistance:

*When there are such large proportions of low-achieving students, it can be more difficult to be successful with Reading Recovery (Smith, 1994).* (Shanahan & Barr, 1995, p. 995)

Because a higher proportion of minority children live in low income areas and minority children are legally protected from educational inequality by the equal protection clause of the U.S. Constitution. **Reading Recovery could potentially violate constitutional law by holding lower expectations for minority children.** At least one site seems to have recognized this as a potential problem:

*In New York City Reading Recovery programs, children are not discontinued until they reach national rather than local averages.* (Shanahan & Barr, 1995, p. 995)

---

## Results

***Will Reading Recovery raise overall school achievement levels?***

**If a school's goal is to raise the overall level of reading performance, Reading Recovery is not the appropriate intervention to choose.** Both Reading Recovery advocates and critics agree on this point. Hiebert (1994) found that Reading Recovery had no positive effects on overall school achievement:

*Despite the implementation of a program with 78,000 students from 1984-1993 in the United States, data from the three primary Reading Recovery sites and from the longitudinal study (DeFord et al., 1990) produce an unconvincing scenario of the effects of Reading Recovery on an age cohort.* (p. 23)

In a response to Hiebert, Reading Recovery promoters, Pinnell and Lyons (1995) agreed that they do not expect Reading Recovery to have an effect on overall achievement:

*"Implementation of the program ... in a given school does not necessarily mean an increase in mean scores but an increase of actual numbers of children at average levels .... [Reading Recovery] will not have lifted the scores of the entire age cohort! It never claimed to do this. Changes in mean scores for the total group may or may not increase; the objective of [Reading Recovery] is to have a larger group of children in the middle range."* (p. 1)

---

***Who actually benefits from Reading Recovery?***

Reading Recovery advocates claim a very high success rate with problem children:

*"Approximately 75-85 percent of the lowest 20 percent of children served by Reading Recovery achieved reading and writing scores in the average range of their class and received no additional supplemental instruction (Pinnell, Lyons, & DeFord, 1988; DeFord, Estice, Fried, Lyons, & Pinnell, 1993; Swartz, Shook, & Hoffman, 1993)."* (Swartz & Klein, 1996).

The independent evaluations find these claims to be exaggerated. First of all, Reading Recovery does not serve all of the lowest children. In a study of Ohio children, the mean national percentile score of children entering Reading Recovery from 1986 to 1991 was not below the 20th percentile, but was 34.5 on the comprehension subscale of the Metropolitan Achievement Test (Hiebert, 1994). Hiebert interpreted this to mean that the **children selected for Reading Recovery come from the 4th quintile (20th to 40th percentile) rather than the bottom quintile (0 to 20th percentile) as claimed.** This higher entry level can be explained in three possible ways: (a) Children are not accepted if they do not meet entry criteria, (b) children who do not make progress are dropped from the program, and (b) a uniform percentage of children in each school are served, regardless of the overall level of performance of children in the school.

***Children not meeting entry level requirements are not accepted.*** Reading Recovery requires that children meet certain criteria on the Clay Diagnostic Measures to enter the Reading Recovery program. In addition, children who are already identified as special education children are not accepted into the program.

Even with these tests and teachers' subjective observations, the potential of children beginning first grade is very difficult to judge. Reading Recovery includes a preprogram phase of "roaming the known" where the Reading Recovery teacher may conduct further informal analyses of the potential of children. During this time the Reading Recovery teacher may reject children who are judged unlikely to benefit from the program. Some of these rejected children are referred to special education. These children are not counted among "served children," because they are considered to have never officially begun Reading Recovery.

Reading Recovery does not report the number of low-performing children who are rejected before lesson 1 because they are not expected to benefit from the program. However, Battelle (1995) included the number of eligible children who were never served in an independent evaluation in Ohio. Eligible children who were never served included the number of rejected children as well as the number of children who never got a turn in Reading Recovery. **Battelle's data indicate that together these children represent 19% of the children originally eligible for Reading Recovery.**

***Children not making progress are dropped.*** The Reading Recovery policy is to anticipate which children will not be successful in Reading Recovery and remove them from the program before completing the full program of 60 lessons. A child who will not benefit from Reading Recovery is to be replaced with a child whom the Reading Recovery teacher believes will benefit, by lesson 12 if possible.

Many of the children who do not seem to benefit are referred to special education. Shanahan and Barr (1995) noted that in Illinois, 7% of the children who began Reading Recovery were referred to special education and, therefore, did not complete the program; and in the Wake County Public School System, 6% were referred to special education. This is in addition to the special education children who were already identified and rejected prior to entry into Reading Recovery. Removing the lowest performing children who make little or no progress in the first several lessons increases the average performance of the remaining group of children. Those children who complete the Reading Recovery program are those children that the Reading Recovery teacher predicted would be able to match the classroom average on the Clay Diagnostic Measures upon completion of the program.

***Reading Recovery can serve only a fraction of the children in first grade.*** The claim to serve the "lowest 20 percent" refers to the local school population, not the national population. When most of the children in a school are in the bottom quintile nationally, Reading Recovery could only serve the lowest 20% of that bottom quintile, leaving 80% of the bottom quintile unserved. On the other hand, when few children are in the bottom quintile, Reading Recovery will serve children performing at higher levels.

Because of the greater availability of Title I funds in low income schools Reading Recovery is more common in those schools.

***How successful is Reading Recovery?***

Children who achieve at the average level of their first-grade class on the book-level measures developed for Reading Recovery are usually non-readers, so success in Reading Recovery rarely means the child is a reader. The average book-level score of a child successfully completing Reading Recovery is level 10 (Shanahan & Barr, 1995). Children scoring at level 10 are not reading authentic text, yet these children would be counted as Reading Recovery successes.

Moreover, the success rates with the children who complete Reading Recovery, as determined by the book-level measures, are exaggerated. Shanahan and Barr (1995) reviewed the national data collected through the Reading Recovery evaluation system (22,193 children) and reported to the National Diffusion Network (DeFord, Estice, Fried, Lyons, & Pinnell, 1993). The success rate was calculated as 84% in the report. However, using the same data and including the 26% who began but did not complete Reading Recovery in the calculation, Shanahan and Barr calculated the percentage of successful children from among all the children served to be only 62%:

*The percentage discontinued [successful] that was reported for the 1991-2 sample, for example, is 84%. Yet, if we were to consider the total number of students served, including those with fewer than 60 lessons, only 62% of the total would be found to complete the program successfully.* (p. 965)

Sixty-two percent is a high estimate because no data were included in the report to the National Diffusion Network regarding the size of the additional group of children who were eligible, but were never served (Shanahan and Barr, 1995). As noted earlier, Battelle's (1995) evaluation for the Ohio Department of Education found that this group represented 19% of the children found eligible for Reading Recovery. If the 19% who were never served are included, the success rate drops to 51%.

This figure for the national data (51%) seems a reasonable estimate because it is very close to Battelle's figure for Ohio (1995). **Battelle found that only 53% of the children eligible for Reading Recovery scored at the classroom average on the book-level measures at the end of first grade.** (19% were not served; 28% were served but did not meet program objectives.):

*Slightly more than 200 students in 36 Ohio school districts were determined to be eligible to receive Reading Recovery services in 1990-91, but did not. Also about 300 of the nearly 875 students who received Reading Recovery services in 1990-91 did not discontinue [were not successful].* (p. 68, Battelle, 1995)

Many reports of Reading Recovery success use the in-house evaluation system and suffer the methodological problems pointed out by Shanahan and Barr (1995). For example, the Southwest Regional Education Board (1996) recently published a report of the Arkansas data gathered through the in-house evaluation system ("Getting Elementary Schools Ready for Children: Reading First"). No independent measures were used to evaluate children's reading performance. Only the Clay Diagnostic measures were used. The report claims an 86% success rate; however, this rate is calculated based on the smaller proportion of children who received a full program. Thirty-one percent of the children served did not receive a full program but these children were not included in the calculation of the success rate. The numbers of children who were eligible but not served was not reported, nor were the numbers of children referred to special education or otherwise dropped from Reading Recovery because of inadequate progress.

In contrast to the high success rates reported for Arkansas from the in-house evaluation system, Pollock (1996), an independent evaluator, found that only 14.7% of the children completing the program reached national norms on standardized measures in the 1995-6 evaluation of Columbus Ohio Public Schools. (This was a higher percentage than Pollock found in the previous years.) This means that only 6.5% of the children originally eligible for Reading Recovery read at grade level at the end of first grade in 1996.

Eligible children represent 20% of the first grade class. To put this in perspective, out of a group of 100 children, only 1 child among the 20 eligible for Reading Recovery would read at grade level due to the Reading Recovery program. Whether other pull-out programs could do as well or better could only be determined with a comparison study.

***Does Reading Recovery do away with the need for other compensatory services?***

Reading Recovery advocates claim that Reading Recovery is so effective that Title I and special education programs are no longer needed, thus creating a savings greater than the expense of Reading Recovery (Dyer, 1992; Southern Regional Education Board, 1996). This argument assumes that by implementing Reading Recovery, children will not need assistance from any other compensatory programs.

The data indicate this assumption is false. As noted earlier, special education children are not accepted into the program and, furthermore, the program removes many children and refers them to special education.

*"It should be noted that Reading Recovery does not do away with early referrals for special education, as the program itself makes many such referrals"* (Shanahan & Barr, 1995, p.987).

Reading Recovery does not do away with the need for Title I services either. Children who successfully complete Reading Recovery often need additional assistance. Pollock (1996) reported that 81% of the children who completed Reading Recovery were still eligible for Title I services in Columbus in the 1995-6 school year. This means that among all children eligible for Title I before Reading Recovery, 92% were still eligible after Reading Recovery. In Columbus over an eight-year period, 31% of the children counted as successful by Reading Recovery standards were still eligible for Title I services (Pollock, 1994).

***Are the effects of Reading Recovery sustained over time?***

Advocates also claim that Reading Recovery enables children to establish a self-extending learning system that allows them to improve as readers after they are returned to the classroom. To support this claim, advocates cite any evidence of growth in reading among discontinued Reading Recovery children without comparing this growth with that of a control group. **Evaluations with a control group find that children who return to the classroom as successfully "recovered" students immediately begin falling behind. Their learning rate is slower than that of other low-achieving children.** For example, Glynn, Crooks, Bethune, Ballard, and Smith (1989) found that the learning rate of Reading Recovery children immediately slowed when the children were returned to the classroom and was much lower than that of other low-achieving children. Data from DeFord, Pinnell, Lyons, and Place (1990) also indicate that Reading Recovery children did not learn at a rate comparable to the average children in the class after being discontinued from Reading Recovery (Shanahan & Barr, 1995). Part of the problem may be that the children have difficulty transferring their strategies of prediction from the predictable text they read in Reading Recovery to the authentic text they might read in the regular classroom.

By third grade, even the effects found on the book-level measures have washed out. *"These results suggest that by third grade, the Reading Recovery instructed groups may not be significantly different from the comparison groups as indicated by measures of text reading"* (Shanahan & Barr, 1995, p.980). Hiebert (1994) found the same pattern of results at fourth grade:

***Although Reading Recovery-tutored students perform better than Achievement Comparison students on an oral reading task, this difference disappears when the task is a standardized one, even one that has the limited passages of the Woodcock Reading Mastery Test--Revised.*** (p. 23)

***Is Reading Recovery more effective than common Title I programs?***

Reading Recovery advocates claim that Reading Recovery is more effective than other reading programs:

*"Studies have shown Reading Recovery to be more effective in achieving short-term and sustained progress in reading and writing than other intervention programs, both one-to-one tutorial and small group methods (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Gregory, Earl, & ODonoghue, 1993)."* (Swartz & Klein, 1996).

The Reading Recovery evaluation system does not compare Reading Recovery with alternative interventions. Separate studies have found that measures using authentic text (standardized measures) generally show no advantage for Reading Recovery over other programs, even at the end of first grade. The Clay Diagnostic measures are generally the only measures that show positive effects for Reading Recovery. For example, Shanahan and Barr (1995) report: *"None of the [early intervention] programs, including Reading Recovery, had any impact on standardized test performance at the end of first grade"* (p. 977). Fincher (1991) compared the performance of children in Reading Recovery with that of children in other compensatory programs in Canton City Schools, Ohio, over a five-year period and found that *"teaching Assistants with almost no training and minimal teaching materials with which to teach and working in less than desirable conditions, outperformed the Reading Recovery teachers when their students' overall achievement was compared."* Fincher also found that when the same teachers teach Reading Recovery and Title I, the teachers get better results with the Title I program.

Wake County Public School System found that Reading Recovery students, *"compared to a control group, were just as likely to be retained, placed in special education, or served in [Title] I a year later"* (WCPPS, 1995, p. ii).

Battelle (1995) also compared Reading Recovery with other compensatory programs independent measures on independent measures. Battelle had a great deal of difficulty getting Ohio schools to administer and submit the results of the independent standardized measures. From the submitted data, Battelle found that at the end of the first year, Reading Recovery students scored only 3.4 percentile points higher than children receiving other common nonindividualized compensatory services in comparison schools. (These services varied considerably from pull-out programs to occasional assistance in the regular classroom.) Even though the small difference was statistically significant, it was not educationally significant according to Battelle.

One frequently cited study (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994) found Reading Recovery more effective than four alternative methods: (a) group delivered Reading Recovery, (b) Reading Recovery delivered by teachers with less training, (c) skills-based instruction without the Reading Recovery framework, and (d) a control. These were the four treatments compared:

1. Reading and Writing was a small-group tutorial program taught by certified teachers trained in Reading Recovery.

2. Reading Success was an individualized tutoring program modeled on Reading Recovery. It was taught by substitute certified teachers who had only some Reading Recovery training.

3. Isolated Skills was individualized tutoring focusing on letters, sounds, words, and text-level strategies. Instruction was based on the classroom basal system. Substitute certified teachers received a 3-day intensive training program and were encouraged to use creativity to explore skills in different ways.

4. The control group was a Title I program using group instruction that focused on practicing skills and learning core words. Teachers received no special training.

Rasinski (1995) points out 3 important methodological flaws in the Pinnell et al. study (1994):

1. Using substitute teachers to teach two of the treatments versus the more experienced teachers who are currently working in the school for the other treatments could have influenced the results.

2. The instructional time varied across treatment conditions, with Reading Recovery children receiving significantly more instruction. Rasinski adjusted the posttest scores by the instructional time factor and found that 5 out of 6 mean scores for children in the comparison treatments were higher than the mean scores of the Reading Recovery group. This means that **when the scores were adjusted to equalize instructional time, the other interventions obtained better results than Reading Recovery.** (See Figure 3 in the Executive Summary.) Only the book-level measures showed consistently better results for Reading Recovery after this adjustment.

3. The comparison of individualized Reading Recovery with small group Reading Recovery did not equalize the teacher time. Two hours of instructional time for teaching 4 students in the individualized format was compared with only 1/2 hour for teaching 4 students in the small group format. **Though the mean scores of both groups were essentially equal, the fact that small group Reading Recovery was four times as efficient as individualized Reading Recovery was overlooked in the report.**

In light of Rasinski's criticisms, many conclusions that Pinnel et al. (1994) made cannot be accepted without further replication. These replications have not occurred.

---

***How does Reading Recovery compare with other research-based interventions?***

None of the interventions compared in the Pinnell et al. study incorporated features of effective instruction to prevent reading disabilities that have recently been identified by the National Institute of Child Health and Human Development program (Center for the Future of Teaching and Learning, 1996). Explicit phonemic awareness instruction in kindergarten with systematic, explicit phonics and extensive practice reading decodable text has had the greatest success in preventing the occurrence of reading problems. Iversen and Tunmer (1993) found that when Reading Recovery was modified to be more systematic, it was 37% more effective. Subsequently, Torgesen, Wagner, Rashotte, and Alexander (in press) found that an explicit systematic approach combined with decodable text was more effective than instruction similar to Iversen and Tunmer's modified Reading Recovery treatment.

Wasik and Slavin (1993) compared the relative effect sizes achieved by five treatments for reading problems in five separate meta-analyses. Two programs had much larger effect sizes than Reading Recovery. Conclusions from such meta-analyses may not be as dependable because of possible differences in the control groups. However, the design of the two more effective programs is consistent with the findings of a significant body of other research, lending support to a conclusion that the two treatments identified as superior by Wasik and Slavin are indeed superior to Reading Recovery. Reading Recovery was not nearly as effective as two programs that provided explicit systematic phonics with extensive practice reading decodable text (the Success for All and Wallach and Wallach programs).

Foorman, Francis, Beeler, Winikates, and Fletcher (in press) found explicit phonemic awareness and phonics combined with decodable text much more effective with Title I children than a whole language program with an unlicensed Reading Recovery support program. The explicit systematic treatment was most effective when it was used in the regular classroom.

*"In order to avoid reading failure, the focus should be on prevention, not intervention. It was the classroom curriculum effect, not the tutorial method effect that was significant. The tutorial effect was not particularly strong, given the weak association between growth in word reading and number of days in tutorial."* (Foorman, Francis, Beerly, Winikates, & Fletcher, in press, p. 16)

These findings support Nicholson's recommendation: *If we ... teach letter-sound correspondences [in the regular classroom], we'll reduce the need for Reading Recovery.* (Nicholson cited in O'Hare, 1995, p. 22)

In fact, many of the instructional techniques used in Reading Recovery are inconsistent with the techniques supported by evidence from scientific intervention research. The findings of the NICHD

research (Center for the Future of Teaching and Learning, 1996) emphasize the value of systematic instruction in phonological skills and the alphabetic principle. The key features of systematic instruction are the following: (a) the lessons are logically organized and planned, and (b) the lessons allow for extensive practice applying phonological skills in decodable text.

Reading Recovery does not provide systematic instruction in letter-sound relationships (the alphabetic principle). Reading Recovery's own document, *Reading Recovery Executive Summary, 1984-1995* describes the unsystematic nature of the instruction in letter-sound relationships:

*"Our approach to phonics does not involve following some prescribed, predefined logical sequence for every student....Many children learn all they need about letter/sound relationships in the process of writing messages, other children are engaged in activities designed to extend their meager knowledge of words"* (p. 8).

Reading Recovery does not use decodable text. **Decodable text is composed of words that use the letter-sound correspondences the children have learned to that point and a limited number of sight words that have been systematically taught.** This allows the children opportunity to practice the letter-sound relationships they have learned in the context of real reading. Reading Recovery uses predictable text which leads the children to use context clues contained in the pictures and provided by the repetitive sentence patterns instead of using sounds for the letters in the word. Predictable text does not give children the opportunity to practice their letter-sound knowledge in the context of real reading. Research shows that overuse of context to figure out unknown words, in fact, "hampers" reading acquisition (Lyon & Chhabra, 1996).

## Costs

### *Is Reading Recovery cost-effective?*

Reading Recovery advocates argue that even when high cost estimates are used, the expense is cheap compared to the multi-year educational costs of special education and Title I and the social costs of letting the child fail to learn to read. Dyer (1992) claimed that although the initial set-up of Reading Recovery is costly, the savings in retentions, Title 1, and Special Education services for districts over the long term is substantial. In fact, Dyer even argued that the short-term annual cost of Reading Recovery is less expensive than first-grade retention.

Dyer's figures for savings (see Table I) assume that (a) Title I and special education are completely ineffective (Hiebert, 1994) and (b) Reading Recovery is always effective and children never make use of Title I or special education services. Both these assumptions are false. Furthermore, Dyer (1992) calculated that the cost per child was only $2063. Dyer's figures omit many costs and assume that one full-time-equivalent teacher works with the recommended number of children and that the success rate is perfect with these children.

Hiebert (1994) and Shanahan & Barr (1995) critiqued Dyer for using hypothetical information and not actual data from the Reading Recovery evaluation system regarding cost and effectiveness. Based on actual effectiveness documented in the Reading Recovery evaluation system, Hiebert and Shanahan and Barr concluded that the cost was much higher per successful child ($4625-$8333). Dyer's figures and the revisions of those cost estimates are summarized below (see Table 1). Niether Hiebert nor Shanahan and Barr obtained figures for all the costs involved, though they noted that these costs include more than teacher salaries. They include the staff development and support for the teacher, the materials the teacher uses, and the set-up of the Reading Recovery teaching area. Shanahan and Barr obtained only some of these additional costs from Reading Recovery sponsors.

A recent California audit (August 6, 1996) from the Joint Legislative Audit Committee found much higher figures for some of the costs than Shanahan and Barr reported from the Reading Recovery sponsors. *"The training for teacher leaders costs approximately $18,300 plus the costs of conferences, travel, and the teacher's salary for the year that they are out of service from their district."* (p. 3) The audit also reported that San Bernardino Unified School District reported per pupil costs of $7000, which

included teacher salaries in the overall cost calculation but did *"not include the $112,000 paid to the foundation at CSU San Bernardino for teacher training"* (pp. 3-4).

Table 1. Dyer's original case for cost effectiveness with Hiebert and Shanahan & Barr's response

| | Dyer, 1992 | Hiebert, 1994 | Shanahan & Barr, 1995 | California audit, 1996 |
|---|---|---|---|---|
| COSTS | | | | |
| Teacher salary | $33,015 | $33,015 | $35,104 (1992-3 ave, AFT) | omitted |
| Benefits | omitted | omitted | $8425 (26% in1992, Bureau of Labor Statistics) | omitted |
| Initial Training per teacher | omitted | omitted | $325* | $2000* |
| Teacher Leader | omitted | omitted | $2042* | $4575* |
| Training room set-up | omitted | omitted | omitted | omitted |
| Yearly conference | omitted | omitted | omitted | omitted |
| Travel | omitted | omitted | omitted | omitted |
| On-going support | omitted | omitted | omitted | omitted |
| Substitutes during training | | | | omitted |
| Instructional materials | | | $350* | $375* |
| Cost per teacher | $33,015 | $33,015 | $46,246 | |
| No. of children served | 16 (the number recommended) | 11 (actual number served) | 8 (number discontinued) | |
| Cost per child | $2063 | $3000 | $4625 | |
| Cost per successful child at end of grade 1 | | $3488 (86% success) | | |
| Cost per successful child at end of grade 4 | | $8333 (36% success) | | |
| SAVINGS per teacher | | | | |
| Retentions | 4 @ $5208 | Dyer's figures | | |
| Chapter 1 children | 4 @ $4715 | assume that all these programs | | |
| Special education placements | 2 @$9906 | are completely ineffective. | | |

* depreciated over 4 years

The Wake County Public School System (WCPSS, 1996) gathered their own data and calculated a cost of $9,211 per successful child:

*The average Reading Recovery teacher serves seven students during a year, and, on average, three or four of those students read at a first-grade level by the end of the year. Annually, the cost per student for all students served in Reading Recovery in WCPSS during 1993-94 was approximately $2,947.50 beyond the regular instructional program. Current evaluation data suggests that by the end of third*

*grade only about two of the students served by a Reading Recovery teacher read at a third-grade level. Thus, the WCPSS has invested approximately $9,211 for each student who is a long-term success.*

*Since the 1990-91 and 1991-92 comparison groups of students who did not receive Reading Recovery achieved a comparable success rate on standardized tests in third grade, and since Reading Recovery expenditures in WCPSS do not seem to have been offset by significant savings from a reduction of need for special education, retention, or Chapter 1 assistance, the program does not appear to be cost effective at this time.*

**Taken together, the data indicate that the cost for Reading Recovery (30 hours of instruction for one child) exceeds the national average per pupil expenditure for one full year of schooling.**

*"We found that Reading Recovery works, but not as well as its proponents claim; that its effects largely dissipate over time; and, that it costs about the equivalent of an additional year of schooling for the children who participate-even accounting for savings in other expenditures"* (Personal communication from T. Shanahan to Assemblyman S. Baldwin, June 26, 1995).

The above cost estimates are in terms of cost per pupil. The cost per school is more useful for discussing how funds can be better allocated to meet the goal of making every child a reader. Assuming a K-5 elementary school population of 600 and a first-grade population of 100, a Reading Recovery program might work with 20 children a year. This requires nearly 4 Reading Recovery teachers or 2 full-time-equivalent teachers. (Each Reading Recovery teacher serves a national average of 5.5 children, Hiebert, 1994.) Of these 20 children, only 1 would be reading at grade level in authentic text by the end of the year (Pollock, 1996). The cost of Reading Recovery in this school would exceed $125,000. (Two teachers cost approximately $100,000 in today's dollars and training and implementation costs exceed $25,000).

These Reading Recovery funds could be reallocated to achieve much greater effectiveness. Foorman et al. (in press) found that the expenditures of huge sums on tutorial programs very similar to Reading Recovery is significantly less cost-effective than implementing effective reading instruction in the classroom. Proper beginning reading instruction in kindergarten and early first grade is not only more effective and less costly but will substantially minimize the number of students needing individual attention.

To accomplish this using the same funds, one of the two FTE Reading Recovery positions (2 of 4 half-time Reading Recovery teachers) would be to used to reduce class size. For example, if there were 4 first-grade classes of 25, adding one more class would reduce the size to 20. The teacher in the remaining Reading Recovery position would be more effective using research-based instruction for children who have reading difficulty. The remaining $25,000 could be invested in a school-wide training program to change classroom instruction to the more effective practices evaluated in the Foorman, et al. study (in press). Based on the data from Foorman et al., a school could expect overall achievement levels to increase more than 1 ½ times by this change alone. Retraining the classroom teachers could be easily accomplished with a one-time investment of $25,000. Retooling a school to use explicit instruction in phonological skills with systematic phonics combined with decodable text is a much more cost-effective alternative. Not only would the lower 20% benefit, but the whole school would benefit from more effective instruction.

## Summary

- **The Reading Recovery data reporting system is flawed.**
- **The standard for successful completion of Reading Recovery is not equitable.**
- **Reading Recovery does not raise overall school achievement levels.**
- **Far fewer students than claimed actually benefit from Reading Recovery.**
- **Children who are not expected to be successful are removed from the program and from the calculation of the success rate.**
- **Reading Recovery does not reduce the need for other compensatory reading services.**
- **Research-based alternative interventions are more effective than Reading Recovery.**

- **Reading Recovery is extremely expensive and does not save other costs.**

---

## REFERENCES


Battelle, Ohio Department of Education. (1995). *Longitudinal Study of Reading Recovery: 1990-91 Through 1993-94.*

Center for the Future of Teaching and Learning. (1996). A synthesis of research on reading from the National Institute of Child Health and Human Development. Http://www.ksagroup.org/center.

Center, Y., Wheldall, K., & Freeman, L. (1992). Evaluating the effectiveness of Reading Recovery: A critique. *Educational Psychology, 12,* 263-273.

Center, Y., Wheldall, K., Freeman, L., Outhred, L., & McNaught, M. (1995). An experimental evaluation of Reading Recovery. *Reading Research Quarterly, 30,* 240-263.

Clay, M. (1993). *Reading recovery, a guidebook for teachers in training.* Portsmouth, NH: Heinemann.

Clay, M. M. ( 1979; 1985). The early detection of reading difficulties. Auckland, NZ: Heinemann.

Clay, M. M., & Cazden, C. B. (1990). A Vygotskian interpretation of Reading Recovery tutoring. In L. Moll (Ed.), Vygotsky and education: Instructional implications and applications of sociohistorical psychology (pp. 206-222). Cambridge, UK: Cambridge University Press.

DeFord, D.E., Estice, R., Fried, M., Lyons, C.E. & Pinnell, G.S. (1993). *The Reading Recovery program: Executive summary 1984-92.* Columbus: The Ohio State University.

DeFord, D.E., Pinnell, G.S., Lyons, C.A., & Place, A.W. (1990). *The Reading Recovery follow-up study, Vol. 11,* 1987-89. Columbus: Ohio State University.

Donley, J., Baenen, N., & Hundley, S. (1993). *A study of the long-term effectiveness of the Reading Recovery Program.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Dyer, P. C. (1992). Reading Recovery: A cost-effectiveness and educational outcomes analysis. Spectrum: Journal of Research in Education, 10(1), 110-119.

Fincher, G. E. (1991). *Reading Recovery and Chapter I: A three-year comparative study.* Canton, OH: Canton City Schools.

Foorman, B., Francis, D., Beeler, T., Winikates, D., & Fletcher, J. (in press). *Early Interventions for Children With Reading Problems: Study Designs and Preliminary Findings.* University of Houston.

Glynn, T., Crooks, T., Bethune, N., Ballard, K., & Smith, J. (1989). Reading Recovery in context: implementation and outcome. *Educational Psychology, 12*(3 & 4), 249-261.

Gregory, D., Earl, L., & ODonoghue, M. (1993). A study of Reading Recovery in Scarborough: 1990-1992. Annual Site Report of the Scarborough School District. Ontario: Scarborough School District.

Groff, P. (1994). Reading Recovery: Educationally sound and cost-effective? *Effective School Practices, 13*(1), 65-69.

Hiebert, Elfrieda. (1994). Reading Recovery in the United States: What Difference Does it Make to an

Age Cohort? *Educational Researcher, 23*(9), 15-25.

Iversen, S., & Tunmer, W. (1993). Phonological Processing Skills and the Reading Recovery Program. *Journal of Educational Psychology, 85*(1), 112-126.

National Diffusion Network. (1993). 1992-93 discontinuation data (Research rep.). Columbus: Reading Recovery National Data Evaluation Center.

Lyon, R., & Chhabra, V. (1996). The current state of science and the future of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews, 1*, 1-8.

Leu, D., DeGroff, L., & Simons, H. (1986). Predictable texts and interactive-compensatory hypotheses: Evaluating individual differences in reading ability, context use, and comprehension. *Journal of Educational Psychology, 78*, 347-352.

Lyons, C.A., Pinnell, G.S., Short, K., & Young, P. (1986). *The Ohio Reading Recovery Project, Vol 4, Pilot Year 1985-86.* Columbus: The Ohio State University.

Nicholson, T. (1991). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology, 83*, 444-450.

Nicholson, T., Lillas, C., & Rzoska, M. (1988). Have we been misled by miscues? *The Reading Teacher, 42*, 6-10.

O'Hare, (1995).

Perfetti, C.A. (1985). *Reading ability*, New York: Oxford University Press.

Pinnell, G. S. (1989). Reading Recovery: Helping at-risk children learn to read. *The Elementary School Journal, 90*(2), 159-181.

Pinnell, G.S., & Lyons, C. (1995). Response to Hiebert: What difference does Reading Recovery make? Unpublished manuscript.

Pinnell, G. S., Lyons, C. A., & DeFord, D. E. (1988). *Reading Recovery: Early intervention for at-risk first graders.* Arlington, VA: Educational Research Service.

Pinnell, G.S., Lyons, C.A., DeFord, D.E., & Bryk, A.S. (1994). Response to Rasinski. . *Reading Research Quarterly, 30*(2), 272-275.

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. Reading Research Quarterly, 29(1), 9-38.

Pollock, J.S. (1994). *Final evaluation report: Reading Recovery program 1995-96.* Columbus, OH: Department of Program Evaluation.

Pollock, J.S. (1996). *Final evaluation report: Reading Recovery program 1993-94.* Columbus, OH: Department of Program Evaluation.

Rasinski, T. (1995). On the effects of Reading Recovery: A response to Pinnell, Lyons, DeFord, Bryk, and Seltzer. *Reading Research Quarterly, 30*(2), 264-270.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood cliffs, NH: Prentice Hall.

Schwartz, R., Moore, P., Schmidt, M., Doyle, M. A., Gaffney, J., & Neal, J. (1996). *Executive Summary of Research on Reading Recovery.*

Shanahan, T. (1987).

Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*(4), 958-996.

Smith-Burke, M. T., Jaggar, A., & Ashdown, J. (1993). New York University Reading Recovery project: 1992 follow-up study of second graders (Research rep.). New York: New York University.

Southern Regional Education Board. (1996). Getting Elementary Schools Ready for Children: Reading First.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 32-71.

Stanovich, K. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education, 5*, 11-19.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Stanovich, K. (1991). Word recognition: Changing perspectives. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 133-180). San Diegor: Academic Press.

Stanovich, K. (1994). Romance versus reality. *The Reading Teacher, 47*(4), 280-291.

Stanovich, K., & Stanovich, P. (1995). How research might inform the debate about early reading acquisition. *Journal of Research in Reading, 18*(2), 87-105.

Stanovich, E., West, R., & Feeman, D. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive-compensatory model. *Journal of Experimental child Psychology, 32*, 185-199.

Swartz, S. L. (1992). Cost comparison of selected intervention programs in California. San Bernardino: California State University.

Swartz, S. L., Shook, R. E., & Hoffman, B. M. (1993). Reading Recovery in California. 1992-93 site report. San Bernardino: California State University.

Swartz, & Klein. (1996). Http:// www.sfusd.k12.ca.us / programs / rr / Rroverview.

Tunmer, W. (1989). *Does Reading Recovery Work?* Massey University.

Wake County Public School System. (1995). Evaluation report: WCPSS Reading Recovery, 1990-94. Raleigh, NC: author.

Wasik, B.A., & Slavin, R.E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly, 28*, 179-200.

West, R., & Stanovich, K. (1978). Automatic contextual facilitation in readers of three ages. *Child Development, 49*, 717-727.